

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 2002, 01:26:26 ; Search time 15.194 Seconds

(without alignments)  
1105.559 Million cell updates/sec

Title: US-09-805-550-2

Perfect score: 2036  
Sequence: 1 MLNWKYTKGNFEIEASPD.....EELTANYLLDHGHEPDDQOO 405

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	632	31.0	409 1	R23B_HUMAN P54727 homo sapien
2	623.5	30.6	416 1	R23B_MOUSE P54728 mus musculu
3	588	28.9	368 1	RH23_SCHPO 074803 schizosacch
4	577	28.3	363 1	R23A_MOUSE P54726 mus musculu
5	573	28.1	363 1	R23A_HUMAN P54725 homo sapien
6	483.5	23.7	398 1	RAD23_YEAST P32628 saccharomyc
7	167	8.2	373 1	DSK2_YEAST P48510 saccharomyc
8	154.5	7.6	354 1	YANG_SCHPO 010169 schizosacch
9	149	7.3	1083 1	T2D3_HUMAN 000268 homo sapien
10	144	7.1	637 1	ODP2_AZOVI P10802 azotobacter
11	140.5	6.9	1509 1	GSRL_HUMAN 094284 homo sapien
12	140	6.9	2441 1	CBP_MOUSE P45481 mus musculu
13	139.5	6.9	1132 1	BAT3_HUMAN P46379 homo sapien
14	137.5	6.8	547 1	ODP2_PSEAE 059658 pseudomonas
15	137.5	6.8	552 1	EPG1_CANAL P43064 candida alb
16	137.5	6.8	865 1	CPN_DROME 002910 drosophila
17	136	6.7	743 1	OCT1_HUMAN P14859 homo sapien
18	136	6.7	797 1	PART_YEAST P25644 saccharomyc
19	135	6.6	352 1	ALGP_PSEAE P15276 pseudomonas
20	135	6.6	1048 1	SR4_RAT 063627 rattus norv
21	134.5	6.6	699 1	VGLG_HSV2H P13280 herpes simp
22	133.5	6.6	1157 1	SR4_HUMAN 095104 homo sapien
23	133.5	6.6	2442 1	CBP_HUMAN 092793 homo sapien
24	131	6.4	1783 1	RAA3_CHIRE 09feca chlamydomon
25	130	6.4	392 1	HME1_HUMAN 005925 homo sapien
26	129.5	6.4	1772 1	MSPI_PLAYO P13828 plasmodium
27	129	6.3	511 1	P60_LISGR 001885 listeria gr
28	129	6.3	688 1	DP3X_HAEIN P43746 haemophilus
29	129	6.3	1794 1	YAVI_SCHPO Q10172 schizosacch
30	127.5	6.3	815 1	RBMS_HUMAN P52756 homo sapien
31	127	6.2	497 1	HMS_DROME P18488 drosophila
32	127	6.2	1544 1	TUSP_HUMAN 09nrt4 homo sapien
33	126.5	6.2	721 1	YK82_MYCTU Q10690 mycobacteri

34	126.5	6.2	825 1	ICP0_HSV2H P28284 herpes simp
35	126.5	6.2	1115 1	NCAL_MOUSE P13595 mus musculu
36	125.5	6.2	1211 1	BUN2_DROME 024592 drosophila
37	125.5	6.2	2090 1	N214_HUMAN P35658 homo sapien
38	125	6.1	2459 1	MABP_RAT P15205 rattus norv
39	124.5	6.1	3149 1	TEGU_EBV P03186 epstein-bar
40	124	6.1	2414 1	P300_HUMAN 009472 homo sapien
41	123.5	6.1	1596 1	MAM_DROME P21519 drosophila
42	123	6.0	339 1	MP5B_LOLPR 040237 lolium pere
43	123	6.0	483 1	PREG_NEUCR 006712 neosporea
44	123	6.0	901 1	A180_MOUSE 061548 mus musculu
45	123	6.0	1326 1	BCC2_ACEXY 082861 acetobacter

## ALIGNMENTS

RESULT 1  
ID R23B\_HUMAN STANDARD; PRT; 409 AA.  
AC P54727;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE UV excision repair protein RAD23 homolog B (HHR23B) (XP-C repair  
DE complementing complex 58 kDa protein) (P58).  
GN RAD23B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=94222030; PubMed=8168482;  
RA Mesutani C., Sugawara K., Yanagisawa J., Sonoyama T., Ue M.,  
RA Enomoto T., Takio K., Tanaka K., van der Spek P.J., Bootsma D.,  
RA Hoeijmakers J.H.J., Hanaoka F.;  
RT "Purification and cloning of a nucleotide excision repair complex  
RT involving the xeroderma pigmentosum group C protein and a human  
RT homologue of yeast RAD23.";  
RT EMBO J. 13:1831-1843(1994).  
RT  
CC -!- FUNCTION: INVOLVED IN DNA EXCISION REPAIR. MAY PLAY A PART IN DNA  
CC DAMAGE RECOGNITION AND/OR IN ALTERING CHROMATIN STRUCTURE TO  
CC ALLOW ACCESS BY DAMAGE-PROCESSING ENZYMES.  
CC -!- SUBUNIT: Heterodimer of a 125 kDa subunit (p125) and of a  
CC 58 kDa subunit (p58). Interacts with MTD.  
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -!- DOMAIN: The ubiquitin-like domain mediates interaction with MTD.  
CC -!- SIMILARITY: CONTAINS 2 UBA DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: D21090; BAA04652.1; -.  
CC HSSP: P54725; IDV0.  
CC Genew: HGNC:9813; RAD23B.  
CC KIM: 6000062; -.  
CC InterPro: IPR004806; Rad23.  
CC InterPro: IPR000449; UBA\_domain.  
CC InterPro: IPR000626; Ubiquitin.  
CC Pfam: PF00240; Ubiquitin; 1.  
CC Pfam: PF00627; UBA; 2.  
CC SMART: SM00165; UBA; 1.  
CC SMART: SM00213; UBO; 1.  
CC TrGFams: TIGR00601; rad23; 1.  
CC PROSITE: PS50053; UBQUITIN\_2; 1.  
CC DNA damage; DNA repair; Nuclear protein.

FT DOMAIN 1 79 UBIQUITIN-LIKE.  
 FT DOMAIN 103 106 POLY-THR.  
 FT DOMAIN 254 260 POLY-ALA.  
 FT DOMAIN 261 269 POLY-THR.  
 FT DOMAIN 336 348 POLY-GLY.  
 SQ SEQUENCE 409 AA; 43171 MW; C026C78273BCB289 CRC64;

Query Match 31.0%; Score 632; DB 1; Length 409;  
 Best Local Similarity 34.1%; Pred. No. 1,1e-28;  
 Matches 150; Conservative 82; Mismatches 138; Indels 70; Gaps 11;

1 MKLNKTLKGTNEFEASPDASVADYKRIETFGOSTYRADQMLIYOGKILKDETTLE 60  
 1 MOVTLTKTLOOQTFKIDIDEETFKALKEIESEKGDAPVAGOKILYAKIINDTALK 60  
 61 SNGVAENSPFLVIMLSKAKASSGASTATTAKAPA-----TLQAPAPVAPASV 109  
 61 EYKIDKKNVYVMTKPKAVSTPAATQOASAPSTTAVTSTTTVAQAPFV-PALAP 119  
 110 ARTPTQ-----APVATAEAPSPVQAPAPAT-VAATD-----ADYISQ 149  
 120 TSTPASITPASATASSEPAPASAKQEKPAEKPAETPVATSPATDSTGSSRSMLFED 179  
 150 AASNLVEGNLEQTLQIILDMGGTWERDVTYVRLAAYNNPERAIDLYSGIPENVEAO 209  
 180 ATSAVTGSGYENMTVEISMKG---YERQVIALALASFNNDRAVEYILMGIPGDRSO 236  
 210 PVARAPAGCOQTNOQASPAOPAVALPVOPSPASAGPNANPLNFPQGVPSGSGSNP 269  
 237 AVVDPPOQA-----STGAPQSSAVAAAAATTAT-----TTSSGHP----- 275  
 270 GAGSGALDLRLQPLQFALLQVQANPOLQPLQELQKONQIILQENQAEFLRLYN 329  
 276 -----LEFLRNQOPQOQRQIILQONPSILPALLOIGRENPQLLOISQHEHFIOLN 329  
 330 E-----SPGCGPGNLTGOLAAAVP---QTLVTPPEERAIORLEGMGRNRELYLEFVA 381  
 330 EYQVQAGGGGGGGGGGGGIAEGSGHMYIQTVPQEKALIERLKALGEPGLIVQATRA 389  
 382 CNKDELTANVLLDHGHEFD 401  
 390 CEKNENIANTFLQONFDED 409

RESULT 2  
 R33B\_MOUSE STANDARD: PRT; 416 AA.  
 ID R33B\_MOUSE  
 AC P54728;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE UV excision repair protein RAD23 homolog B (MHR23B) (XP-C repair  
 DE complementing complex 58 kDa protein) (P58).  
 GN RAD23B OR MHR23B.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/c; TISSUE-Testis;  
 RX MEDLINE=96403997; PubMed=8808275;  
 RA van der Spek P.J., Visser C.E., Hanacka F., Smit B.,  
 RA Hagemeijer A., Bootsma D., Hoeljmakers J.H.J.;  
 RT "Cloning, comparative mapping, and RNA expression of the mouse  
 RT homologues of the Saccharomyces cerevisiae nucleotide excision repair  
 RT gene RAD23."  
 RL Genomics 31:20-27(1996).  
 CC -1- FUNCTION: INVOLVED IN DNA EXCISION REPAIR. MAY PLAY A PART IN DNA  
 CC DAMAGE RECOGNITION AND/OR IN ALTERING CHROMATIN STRUCTURE TO  
 CC ALLOW ACCESS BY DAMAGE-PROCESSING ENZYMES.  
 CC -1- SUBUNIT: HETERODIMER OF A 125 KDA SUBUNIT (P125) AND OF A  
 CC 58 KDA SUBUNIT (P58).

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 UBI DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC -----

DR EMBL: X92411; CA63146.1; -  
 DR HSSP: P54725; IDV0.  
 DR MGD: MGI:105128; Rad23b.  
 DR InterPro: IPR004806; Rad23.  
 DR InterPro: IPR004409; UBA.domain.  
 DR InterPro: IPR006626; Ubiquitin.  
 DR Pfam: PF00627; UBA; 2.  
 DR SMART: SM00165; UBA; 2.  
 DR SMART: SM00213; UBO; 1.  
 DR TIGRFAMs: TIGR00601; rad23; 1.  
 DR PROSITE: PS50053; UBIQUITIN\_2; 1.  
 DR DNA damage; DNA repair; Nuclear protein.  
 FT DOMAIN 1 79 UBIQUITIN-LIKE.  
 FT DOMAIN 255 261 POLY-ALA.  
 FT DOMAIN 262 270 POLY-THR.  
 FT DOMAIN 336 355 POLY-GLY.  
 SQ SEQUENCE 416 AA; 43516 MW; 13E0245AD892205 CRC64;

Query Match 30.6%; Score 623.5; DB 1; Length 416;  
 Best Local Similarity 32.4%; Pred. No. 3.3e-28;  
 Matches 146; Conservative 85; Mismatches 136; Indels 83; Gaps 10;

1 MKLNKTLKGTNEFEASPDASVADYKRIETFGOSTYRADQMLIYOGKILKDETTLE 60  
 1 MOVTLTKTLOOQTFKIDIDEETFKALKEIESEKGDAPVAGOKILYAKIINDTALK 60  
 61 SNGVAENSPFLVIMLSKAKA-----SSGASTATTAKAPATLQAPAPVAPASV 110  
 61 EYKIDKKNVYVMTKPKAVTAVPATTPQSSPTPTVSSPAPV-----AAQAPAPTPA 116  
 111 RTPTQAPVAT-----AETAP-----PSVOPQAPAPATVATD-----ADV 146  
 117 LAPSTPASTTPASTTASSEPAPACATQEPKPAEKPAQTPVLTSPAPADSPFGDSSRNL 176  
 147 YQASNLVEGNLEQTLQIILDMGGTWERDVTYVRLAAYNNPERAIDLYSGIPENY 206  
 177 FEDATSAVTGSGYENMTVEISMKG---YERQVIALALASFNNDRAVEYILMGIPGDR 233  
 207 EAPVAPAPAGCOQTNOQASPAOPAVALPVOPSPASAGPNANPLNFPQGVPSGSGSNP 266  
 234 ESQAVVDPPOQAVST-----GTQSPAVAAAAATTAT-----TTSSGHP- 275  
 267 VVPGAGSGALDLRLQPLQFALLQVQANPOLQPLQELQKONQIILQENQAEFLRLYN 326  
 276 -----LEFLRNQOPQOQRQIILQONPSILPALLOIGRENPQLLOISQHEHFIQ 326  
 327 LVNESP-----GPGGNLTGOLAAAVPQTLVTPPEERAIORLEGMGRN 371  
 327 MLNEPVQAGGGGGGGGGGGGIAEGSGHMYIQTVPQEKALIERLKALGEP 386  
 372 RELVLEFPACNKNDELTAANYLLDHGHEFD 401  
 387 EGVLTQATPACKNENIANTFLQONFDED 416

RESULT 3  
 RH23\_SCHPO STANDARD: PRT; 368 AA.  
 ID RH23\_SCHPO  
 AC 074803;  
 DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE UV excision repair protein rhp23 (RAD23 homolog).  
 GN RHP23 OR SPBC2D10.12.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 ON NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.  
 RC STRAIN=SP223;  
 RX PubMed=11786722;  
 RA Elder R.T., Song X.-Q., Chen M., Hopkins K.M., Lieberman H.B.,  
 RA Zhao Y.;  
 RT "Involvement of rhp23, a Schizosaccharomyces pombe homolog of the  
 RT human HHR23A and Saccharomyces cerevisiae Rad23 nucleotide excision  
 RT repair genes, in cell cycle control and protein ubiquitination.";  
 RL Nucleic Acids Res. 30:581-591(2002).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Stevens K.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Sharp S.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,  
 RA Weytens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Beger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,  
 RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Kocher M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 RL [3]  
 RN CHARACTERIZATION.  
 RP PubMed=10652237;  
 RX Lombers M., Goeloe J.I., den Dulk H., Brandsma J.A., Brouwer J.;  
 RT "Identification and characterization of the rhp23(+) DNA repair gene  
 RT in Schizosaccharomyces pombe.";  
 RL Blochm. Biophys. Res. Commun. 268:210-215(2000).  
 CC -1- FUNCTION: Involved in postreplication repair of UV-damaged DNA.  
 CC Postreplication repair functions in gap-filling of a daughter  
 CC strand on replication of damaged DNA.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF174293; AAD51975.1; -;  
 CC DR EMBL: AL031788; CAA21170.1; -;  
 CC DR HSP: P34725; IDV0.

DR InterPro: IPR000449; UBA.  
 DR InterPro: IPR000626; Ubiquitin.  
 DR Pfam: PF00627; UBA; 2.  
 DR Pfam: PF00240; ubiquitin; 1.  
 DR SMART: SM00165; UBA; 2.  
 DR SMART: SM00213; UBA; 1.  
 DR PROSITE: PS50053; UBIQUITIN\_2; 1.  
 DR DNA damage; DNA repair; Nuclear protein.  
 KW DNA damage; DNA repair; Ubiquitin-like.  
 FT DOMAIN 1 77  
 FT DOMAIN 119 122  
 FT DOMAIN 205 208  
 FT POLY-ALA.  
 FT POLY-GLN.  
 SQ SEQUENCE 368 AA; 40135 MW; 5CE75EB7E190EFD4 CRC64;  
 Query Match 28.9%; Score 588; DB 1; Length 368;  
 Best Local Similarity 35.3%; Pred. No. 2.8e-26;  
 Matches 146; Conservative 72; Mismatches 138; Indels 58; Gaps 13;  
 QY 1 MKNVKTLLKGTNEET-EASPDASVADYKRIETGOGSTYRADQOMLTIGYKILKDETL 59  
 DB 1 MNLTFKMLQOQKEVYISVDSADTKISELKERIQT---QONVEVEROKLYSGRIADDKTV 57  
 QY 60 ESNQVANSFLVIMLSKAKSSGASATATKAPATLAOPAPVAPASVARTPTQAPVA 119  
 DB 58 GEYIKQEDTIVCMVSRPKTSTS-----TPKSASRPAP-NPASPVEKKVEAPSS 106  
 QY 120 TAETAPSVQPOAPAPAT---VAATDDADYYSQAASMLVGNMLEQTIQOILDMGSGTWE 176  
 DB 107 TVASTSTTQGVAAASNDPTATSEAPI---DANTLAVGAQKNVAVENMVEG---YE 160  
 QY 177 RDIYVRLRAAYNPPERPAIDLYSGIPBNV---EAQPVAPAPAGQOTNOAASPAPAV 233  
 DB 161 RSEVERMRAPAFNPNDAVEVILLGIPEDILNRQRESAALAAQOQOSALA----- 213  
 QY 234 ALPQSPASAGPAPANPLNLFPGVPSGSGNPVGPV-AGSGALDALRQLPQALLQIV 292  
 DB 214 -----PTSG---QPANLFQALSENENQEPSTVGGDDPLGTRSLIPQQLQIV 263  
 QY 293 QANPQILQPMLELQKGNQPNQILRIQENQAEFLRLVNESPEGG---PGNITGLQLAAPV 349  
 DB 264 QONQPMLETLIQGGQDPAQAQIQTNPFAFIQLMEAGEGESALPSGGI----- 314  
 QY 350 QTLVTPEERPAIORLESGMFRRLVYEVFPACKDELTANLYLDHGHEFDDQ 403  
 DB 315 -QIDITQEESESIDRLCOLGPDNIVIVAYLADCKNEELNAYLFEHGHESEDE 367  
 RESULT 4  
 R23A\_MOUSE  
 ID R23A\_MOUSE STANDARD; PRT; 363 AA.  
 AC P54726;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE UV excision repair protein Rad23 homolog A (HHR23A).  
 GN RAD23A OR MHR23A.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C; TISSUE=Testis;  
 RX MEDLINE=96403997; PubMed=808275;  
 RA van der Spek P.J., Visser C.E., Hanaoka F., Smit B.,  
 RA Hagemeljer A., Bootsma D., Hoeijmakers J.H.J.;  
 RT "Cloning, comparative mapping, and RNA expression of the mouse  
 RT homologues of the Saccharomyces cerevisiae nucleotide excision repair  
 RT gene RAD23.";  
 RL Genomics 31:20-27(1996).  
 CC -1- FUNCTION: Involved in postreplication repair of UV-damaged DNA.  
 CC Postreplication repair functions in gap-filling of a daughter  
 CC strand on replication of damaged DNA (probable).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (probable).

CC -1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 2 UBA DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: X92410; CAA63145.1; -  
DR HSSP: P54725; IDV0  
DR MGD: MGI:105126; Rad23a.  
DR InterPro: IPR004806; Rad23.  
DR InterPro: IPR000449; UBA\_domain.  
DR InterPro: IPR000626; Ubiquitin.  
DR Pfam: PF00240; ubiquitin; 1.  
DR Pfam: PF00627; UBA; 2.  
DR SMART: SM00165; UBA; 2.  
DR SMART: SM00213; UBO; 1.  
DR TIGRFSMS: TIGR00601; rad23; 1.  
DR PROSITE: PS50053; UBIQUITIN\_2; 1.  
DR DNA damage: DNA repair; Nuclear protein.  
KM DOMAIN 1 79 UBIQUITIN-LIKE  
FT SEQUENCE 363 AA; 39769 MW; 67EAB96EBAB5C203 CRC64;  
SQ  
Query Match 28.3%; Score 577; DB 1; Length 363;  
Best Local Similarity 33.4%; Pred. No. 1.9e-25;  
Matches 140; Conservative 77; Mismatches 124; Indels 78; Gaps 13;  
QY 3 LNKVTLKGFNFELASPDASVADVKRIETTGQSTYRADQQLIYOGKILKDETTLESN 62  
DB 5 IRLKTLQOQTFKIRMBDETVAVLKEKIEAKGRAPFVAGOKILYAGKILSDVPIDRY 64  
QY 63 GVAENSLVYIMSKAKASSGASTATTAQAPATLAQAPAVP--AASVARRP----- 113  
DB 65 HIDEKRFVVMYTKAKA--GOGISAPPEA--SPTRAVPEPSTPPPVVLASGSHPPYSREK 122  
QY 114 TQAPVATATAPSPVQQAAPATVATDADVYSQAASNLVFGNNLEQTIOOILDMGCG 173  
DB 123 SPSEESTYTTSPESIGSVSSGSGRED-----AASLTVTGSEYETMLTEIMSG-- 174  
QY 174 THERDVTVAALRAAYNNPERALDLYSGTPENVEAPVARAPAGQ--QTNQQAASPAQDA 232  
DB 175 -TEREVVAVALRASNNPRAVEYLLTGIGSPE-----PEHGSVQDSORAEOPATEA 226  
QY 223 VALPVQSPASAGPNANPLNLFPGQVPSGSGNPGVAGSGALDLALROLPOFALLOLV 292  
DB 227 -----AGENPLEFLRQPOFOFNKROYI 248  
QY 293 QANPOLIQPMLOELGKONPOLIRLQENQAEFLRLVNESPEGGPGNT-----LGQLAA 346  
DB 249 QONPALPALLOOLGQENPOLLOQISRHQEOFIQMLNEPP--GELADISDEGEVGAIGE 306  
QY 347 AVPQOT--LVTPEEREALIORLGEMGNRLVLEVFPACKKDEELTANYLLDGHCHDDQ 403  
DB 307 EAPQNNYIOVTFQEKAEIRLKAIGPESLVIOAYFACENENLANFL--SQNPDE 363  
RESULT 5  
R23A\_HUMAN  
ID R23A\_HUMAN STANDARD; PRT; 363 AA.  
AC P54725;  
DT 01-OCT-1996 (Rel. 34; Created)  
DT 01-OCT-1996 (Rel. 34; Last sequence update)  
DT 15-JUN-2002 (Rel. 41; Last annotation update)  
DE UV excision repair protein RAD23 homolog A (HHR23A).  
GN RAD23A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-9422030; PubMed-8168482;  
RA Masutani C., Sugawara K., Yanagisawa J., Sonoyama T., Ui M.,  
RA Enomoto T., Takio K., Tanaka K., van der Spek P.J., Bootsma D.,  
RA Hoeijmakers J.H.J., Hanawalt F.,  
RT Purification and cloning of a nucleotide excision repair complex  
RT involving the xeroderma pigmentosum group C protein and a human  
RT homologue of yeast RAD23.  
RL EMBO J. 13:1831-1843(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Lamerdin J., McCreedy P., Stillwagen S., Ramirez M., Carrano A.;  
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP STRUCTURE BY NMR OF 319-363.  
RX MEDLINE-99061330; PubMed-9846873;  
RA Dieckmann T., Withers-Ward E.S., Jarosinski M.A., Liu C.F.,  
RA Chen I.S.Y., Feigon J.;  
RT Structure of a human DNA repair protein UBA domain that interacts  
RT with HIV-1 Vpr.  
RL Nat. Struct. Biol. 5:1042-1047(1998).  
RN [4]  
RP STRUCTURE BY NMR OF 319-363.  
RX MEDLINE-20541363; PubMed-11087358;  
RA Withers-Ward E.S., Mueller T.D., Chen I.S.Y., Feigon J.;  
RT Biochemical and structural analysis of the interaction between the  
RT UBA(2) domain of the DNA repair protein HHR23A and HIV-1 Vpr.  
RL Biochemistry 39:14103-14112(2000).  
CC -1- FUNCTION: Involved in postreplication repair of UV-damaged DNA.  
CC Postreplication repair functions in gap-filling of a daughter  
CC strand on replication of damaged DNA (Potential).  
CC -1- SUBUNIT: Interacts with MJD.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- DOMAIN: The ubiquitin-like domain mediates interaction with MJD.  
CC -1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 2 UBA DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: D21235; BAA04767.1; -  
DR EMBL: AD000092; AAB51177.1; -  
DR PDB: IDV0; 11-FEB-00.  
DR GENE: HGNC:9612; RAD23A.  
DR MIM: 600061; -  
DR InterPro: IPR004806; Rad23.  
DR InterPro: IPR000449; UBA\_domain.  
DR InterPro: IPR000626; Ubiquitin.  
DR Pfam: PF00240; ubiquitin; 1.  
DR Pfam: PF00627; UBA; 2.  
DR SMART: SM00165; UBA; 2.  
DR SMART: SM00213; UBO; 1.  
DR TIGRFSMS: TIGR00601; rad23; 1.  
DR PROSITE: PS50053; UBIQUITIN\_2; 1.  
DR DNA damage: DNA repair; Nuclear protein; 3D-structure.  
KM DOMAIN 1 81 UBIQUITIN-LIKE  
FT SEQUENCE 363 AA; 39609 MW; C4E47E9313B84785 CRC64;  
SQ  
Query Match 28.1%; Score 573; DB 1; Length 363;  
Best Local Similarity 33.3%; Pred. No. 1.9e-25;  
Matches 139; Conservative 79; Mismatches 124; Indels 76; Gaps 13;  
QY 3 LNKVTLKGFNFELASPDASVADVKRIETTGQSTYRADQQLIYOGKILKDETTLESN 62  
DB 5 IRLKTLQOQTFKIRMBDETVAVLKEKIEAKGRAPFVAGOKILYAGKILSDVPIDRY 64  
QY 63 GVAENSLVYIMSKAKASSGASTATTAQAPATLAQAPAA--VPAPASVARTPTQAPVA 119

```
Db 65 RIDEKNEVWVTKTKA---GQGTSAPEASPTIAPSSSTFPAPYSGMSHP---PAA 118
QY 120 TAETAP-----PSYQOAPAAATVAATDDADVYSOASNLVFGNNLTQITQIILDMGCGT 174
Db 119 REDKSPSEASPTTSPESV-SGSVPSSSGSGREDASTLVTSYEYETMLTEISMG--- 174
QY 175 WERTVYRALRAAYNNPERAIDYISGIPENVEAQVAPAPACQ-OTNOQASPAQAPV 233
Db 175 YERERVYAAALRASYNHRAVEYLLTGITPGSP-----PEHSGVQSEQSEAPATEA- 226
QY 234 ALPVQSPASAGPRANLNFPGVSPGSGSNPGVVPAGSGALDALQIQFOQALLQYQ 293
Db 227 -----AGENPLEFLKDOQPOFQNMKROYIQ 249
QY 294 ANFOITLQPMLOELQKONPOLILRLIQENQAFELRLVNESPEGPGCNI-----LGOATAA 347
Db 250 QNALLPALLQOLQOGEQPOLLOQSRHQEQFIQMLNEPP--GLADISDVEGEGVAGIEE 307
QY 348 VPQT--LTVPREERAIQRLQEGMGNELVLEVFACNKEDELTANYLIDHGHEFDQ 403
Db 308 APOMNVIQVTPQEKALERKALGFPESLVIQAFACEKNENIAPLTL--SQNFDBE 363
```

## RESULT 6

```
RA23_YEAST STANDARD; PRT; 398 AA.
AC P32628;
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-1993 (Rel. 27, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN RAD23 OR YEL037C OR SYGP-ORF29.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID:4932;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-94067136; PubMed-8246991;
RA Watkins J.F., Sung P., Prakash L., Prakash S.;
RT "The Saccharomyces cerevisiae DNA repair gene RAD23 encodes a nuclear
RT protein containing a ubiquitin-like domain required for biological
RT function."
RL Mol. Cell. Biol. 13:7757-7765(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
RA Mei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288c / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berro A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Myskale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Mei Y., Yellon M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-B-6441;
RX MEDLINE-94016558; PubMed-841151;
RA Melnick L., Sherman F.;
RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
RT of Saccharomyces cerevisiae share a common ancestry."
RL J. Mol. Biol. 233:372-388(1993).
CC - SUBCELLULAR LOCATION: Nuclear (Probable).
CC - SIMILARITY: CONTAINS 1 UBQUITIN-LIKE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC EMBL; L25428; AAA16070.1; -
DR EMBL; U18779; AAB65005.1; -
DR EMBL; L22172; AAA34935.1; -
DR EMBL; L22173; AAA34938.1; -
DR EMBL; S65964; AAD13972.1; -
DR EMBL; S66117; AAB28441.1; -
DR PIR; S30845; S30845.
DR HSP; P54725; IDV0.
DR SGD; S0000763; RAD23.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR000449; UBA_domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBO; 1.
DR TIGRfam; TIGR00601; rad23; 1.
DR PROSITE; PS50053; UBQUITIN_2; 1.
KW DNA damage; DNA repair; Nuclear protein.
FT DOMAIN 1 77 UBQUITIN-LIKE.
FT CONFICT 277 277 A -> R (IN REF. 4).
SQ SEQUENCE 398 AA; 42366 MW; B3F0436DAB6B833 CRC64;
```

Query Match 23.78; Score 483.5; DB 1; Length 398;  
Best Local Similarity 32.38; Pred. No. 2e-20;  
Matches 142; Conservative 65; Mismatches 144; Indels 89; Gaps 18;

```
QY 1 MKLWKLTKGNFELRSPDASVADVKRIETQGSYRADQOM-LIYQKTKKDETL 59
Db 2 VSLTFKFKKKKVPDLPSWTILETK---TKLAQSTSCSESIKILSKVLDQSKTV 57
QY 60 ESNQVAENSLVIMLSKAKSSSGASTATYAKAPATLAQAPAPAPASVAKRP----- 113
Db 58 SEGKJQDDQYVFNVSQK-----STKTKYTP-----PLAPES--ATTGRENST 101
QY 114 -----TOAPVATTAETAPSPVOPQAPATVATDDADVYSOASNLVFGNNLTQITQIIL 168
Db 102 EASPTDASAPATAEAGSQPOEQTATERTESA-----STPGFVGTERTETIRIM 156
QY 169 DMGGTWERDITVYRALRAAYNNPERAIDYISGIPENVEAQVAPAPACQOTNOQASP 228
Db 157 EMG---TQREVERALRAAFNNPRAVEYLLMGIPENL-----RQPPQQT---AAA 204
QY 229 AQPVALVPQSPASAGPRANLNLFPQVSPGSGSNPGVVPAGSGALDALQRLP----- 283
Db 205 EQPTAATTAATPAED-----DLFQAAGGASASAL-GTTCGATDDAAGGPPSGTG 256
QY 284 -----QFOALLQVQANFOITLQELQKONPOLILRLIQENQAFELRLVNES----- 331
Db 257 LTVEDLILRQVVGNGPEALPLLENISARYPQIREHMANPEFVSLLEAVQDNQDV 316
QY 332 PEG-----GGGNILGOLAA-VPO-----LTVPREERAIQRLQEGMGNELVLE 377
Db 317 MEGADMYEGEDIEVTGAAAGAGQEGEGSFQVDYTPEDDAISRICEIGFERFDLVIQ 376
QY 378 VFACNKEDELTANYL-LD 396
Db 377 VYFACNKEEAANILPSDH 396
```

## RESULT 7

```
DSK2_YEAST STANDARD; PRT; 373 AA.
AC P48510;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
```

DT 16-OCT-2001 (rel. 40, last annotation update)  
 DE Ubiquitin-like protein DSK2.  
 GN DSK2 OR SHE4 OR YMR276W OR YMR021.02.  
 OS Saccharomyces cerevisiae (baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE=96281973; PubMed=8682868;  
 RA Biggins S., Ivanovska I., Rose M.D.;  
 RT "Yeast ubiquitin-like genes are involved in duplication of the  
 RT microtubule organizing center";  
 RL J. Cell Biol. 133:1331-1346(1996).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RX Pearson D., Bowman S., Barrell B.G., Rajandream M.A.;  
 RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INVOLVED, WITH RAD23 IN SPINDLE POLE BODY DUPLICATION.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: To S.POMBE SPAC26A3.16.  
 CC -1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L40587; AAB07267.1;  
 DR EMBL: 249704; CAA89774.1;  
 DR HSSP: 015843; INDD.  
 DR SGD: S0004889; DSK2.  
 DR InterPro: IPR000449; UBA\_domain.  
 DR InterPro: IPR000526; Ubiquitin.  
 DR Pfam: PF00240; Ubiquitin.1.  
 DR Pfam: PF00627; UBA.1.  
 DR SMART: SM00165; UBA.1.  
 DR SMART: SM00213; UBO.1.  
 DR PROSITE: PS00299; UBIQUITIN\_1; 1.  
 DR PROSITE: PS50053; UBIQUITIN\_2; 1.  
 DR Nuclear protein.  
 KW DOMAIN 1 76 UBIQUITIN-LIKE.  
 FT FT 109 109 R -> A (IN REF. 2).  
 FT CONFLICT 296 296 R -> A (IN REF. 2).  
 FT SEQUENCE 373 AA; 39516 MW; 25DF82B9DB67DF6 CRC64;  
 SQ  
 Query Match 8.2%; Score 167; DB 1; Length 373;  
 Best Local Similarity 23.2%; Pred. No. 0.0081;  
 Matches 94; Conservative 47; Mismatches 168; Indels 96; Gaps 19;  
 QY 1 MLNWKTKLGN-FEIEASPDASVADVKRIETFGQSTYRADQMLTYOGILNDETL 59  
 DB 1 MSINHIKSGDKWEVNAPESTVLOFKKAIKANG--IPVANOGLISGKILKDKQIV 57  
 QY 60 EENGVAESEFLVIMSKAK--ASSSGASTATTAKAPATLAPPAVAAVAARPTQA 116  
 DB 58 ESYHIQDGHSHLVKSQPPQOTASAGANNATATGA--AACTGATPNN 104  
 QY 117 PVATAETAPSPVQPAAPATVATDDADYVSOAASNLVFGNNLEPOTIQOILDMGGTWE 176  
 DB 105 S-GQGRGFRLADLT SARAGYLNPMSADMFPGDGL--NDSNNDGLLRM---ME 156  
 QY 177 RDTVYRALFAAYNPERAIDYISGIPENVEAQPVARAPDAQQTNQAAAP--AQA 232  
 DB 157 NFIFSONNEMSLNPQ-MIDFMIQSNPQLQAMGPQARQMLQSPMFRMLTNDPIRQSQ 215  
 QY 223 VALPYQSPASAGPANPLNLFPGGVPS-----GSPGVVPAAGSQA 275

DB 216 FARMDPN-AGMGSAGGAASAPP--APGGDAPEEGSNTNTSSNTGNAGTNA 272  
 QY 276 LDALROLPOFOALLQVQANP--QLIQPMLOELGKONPOLILIOBNAEFLRLNESPE 333  
 DB 273 NFA-----ANFASLINPL-----NP-----FANAGNRAST 299  
 QY 334 GPGGN--ILGOLAAVPTL--TVTPEE--EAIQRLGEMGF 370  
 DB 300 GMFAEDPALLASMFQPPVQASQAEPTRPPEEREYHQLRLNDMGF 344  
 RESULT 8  
 YADG\_SCHPO STANDARD; PRI; 354 AA.  
 ID YADG\_SCHPO  
 AC Q10169;  
 DT 01-OCT-1996 (rel. 34, created)  
 DT 01-OCT-1996 (rel. 34, last sequence update)  
 DT 15-JUN-2002 (rel. 41, last annotation update)  
 DE Hypothetical protein C26A3.16 in chromosome I.  
 GN SPAC26A3.16.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomycetes.  
 NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 DR Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 DR Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 DR Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 DR Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 DR Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 DR Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,  
 DR James K., Jones L., Jones M., Leather S., McDonald D., McLean J.,  
 DR Mooney P., Moule S., Mungall K., Murphy L., Nidlett D., Odell C.,  
 DR Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 DR Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 DR Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 DR Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 DR Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,  
 DR Wellens I., Vansireels E., Rieger M., Schefer M., Mueller-Auer S.,  
 DR Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
 DR Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 DR Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 DR Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,  
 DR Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 DR Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 DR Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 DR Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
 RA Shpakovskii G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe";  
 RL Nature 415:871-880(2002).  
 CC -1- SIMILARITY: TO YEAST DSK2.  
 CC -1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z69240; CA93239.1;  
 DR InterPro: IPR000449; UBA\_domain.  
 DR InterPro: IPR000526; Ubiquitin.  
 DR Pfam: PF00240; Ubiquitin.1.  
 DR Pfam: PF00627; UBA.1.  
 DR SMART: SM00165; UBA.1.  
 DR SMART: SM00213; UBO.1.



```

Db 214 SLVNNGBA-----ALPLPKPAAGTVIQTFFPGAAAPADAA5PPAAPAPAA 267
Qy 113 PTOAPVATETAPPVQPOQAPAPATVAATDDADVY5QAASNLVFNLEQITQOIIDMG 172
Db 268 PPPPPAPATLAPRPGHPAPPTAPAPVPPA-----AQN-----G 304
Qy 173 GTWERPTVVRAALRAAYNNPERAIDYLSGIPENEVAP-----VARAPAAQOQTNQ----- 223
Db 305 GS-----AGAAPAPAPFA-----AGGPAGVSCGPGGAAAAAPAPGVKAESPRTV 349
Qy 224 QAASPAQAPVALVPPSPASA-----GPNANPLNFPQGV--PSCGSNPGVYPGASGALDA 278
Db 350 QAAPPPAAQTLAASGPASTASWITGPTMOGALFSPAAPVPPAPGTPTGLPKGAAGAVTOS 409
Qy 279 LRLQLPQ-----FQALLQLVQANPOLIQLPMLDELKONPOI---LRLIOENAEFLRYNE 330
Db 410 LSTPTATITSGIRATLTPYTLARLPQPPQNPNTINQNFQLPFGMVLYRSNGULMIPOQ 469
Qy 331 SPEGPGCNILGOLAA---AVPQTLFVTP 356
Db 470 A-----LAQMOQAHAHQPT-TMAP 488

```

```

RESULT 10
ODP2_AZOV1
ID ODP2_AZOV1 STANDARD; PRT; 637 AA.
AC P10802;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
DE complex (EC 2.3.1.12) (E2).
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 478;
RX MEDLINE=88271330; PubMed=3292237;
RA Hanemaaijer R., Janssen A., de Kok A., Veeger C.;
RT "the dihydrolipoalyltransacetylase component of the pyruvate
RT dehydrogenase complex from Azotobacter vinelandii. Molecular cloning
RT and sequence analysis."
RL Eur. J. Biochem. 174:593-599(1988).
RN [2]
RP SEQUENCE OF 1-15 AND 380-415.
RX MEDLINE=88082750; PubMed=3691494;
RA Hanemaaijer R., de Kok A., Jolles J., Veeger C.;
RT "The domain structure of the dihydrolipoalyl transacetylase component
RT of the pyruvate dehydrogenase complex from Azotobacter vinelandii."
RL Eur. J. Biochem. 169:245-252(1987).
RN [3]
RP LIPOYL DOMAIN CONFORMATION.
RX MEDLINE=89052887; PubMed=3191993;
RA Hanemaaijer R., Vervoort J., Westphal A.H., de Kok A., Veeger C.;
RT "Mobile sequences in the pyruvate dehydrogenase complex, the E2
RT component, the catalytic domain and the 2-oxoglutarate dehydrogenase
RT complex of Azotobacter vinelandii, as detected by 600 MHz 1H-NMR
RT spectroscopy."
RL FEBS Lett. 240:205-210(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 381-637.
RX MEDLINE=92196586; PubMed=1549782;
RA Matvevi A., Obojlova G., Schulze E., Kalk K.H., Westphal A.H.,
RA de Kok A., Hol W.G.J.;
RT "Atomic structure of the cubic core of the pyruvate dehydrogenase
RT multienzyme complex."
RL Science 255:1544-1550(1992).
RN [5]
RP STRUCTURE BY NMR OF 1-78.
RX MEDLINE=94422112; PubMed=8068086;

```

```

RA Berg A., de Kok A., Vervoort J.;
RT "Sequential 1H and 15N nuclear magnetic resonance assignments and
RT secondary structure of the N-terminal lipoyl domain of the
RT dihydrolipoalyl transacetylase component of the pyruvate dehydrogenase
RT complex from Azotobacter vinelandii."
RL Eur. J. Biochem. 221:87-100(1994).
RN [6]
RP STRUCTURE BY NMR OF 1-78.
RX MEDLINE=97234563; PubMed=9119000;
RA Berg A., Vervoort J., de Kok A.;
RT "Three-dimensional structure in solution of the N-terminal lipoyl
RT domain of the pyruvate dehydrogenase complex from Azotobacter
RT vinelandii."
RL Eur. J. Biochem. 244:352-360(1997).
CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
CC multiple copies of three enzymatic components: pyruvate
CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
CC lipoamide dehydrogenase (E3).
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-
CC acetyldihydrolipoamide.
CC -1- COFACTOR: THE E2 COMPONENT CONTAINS THREE COVALENTLY-BOUND LIPOYL
CC COFACTORS.
CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
CC SYMMETRY.
CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
CC -1- SIMILARITY: CONTAINS 3 LIPOYL-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X12455; CAA30987.1; ALT_INIT.
DR PIR: S01017; XXAV.
DR PDB: 1EAA; 31-OCT-93.
DR PDB: 1EAB; 31-OCT-93.
DR PDB: 1EAC; 31-OCT-93.
DR PDB: 1EAD; 31-OCT-93.
DR PDB: 1EAE; 31-OCT-93.
DR PDB: 1EAF; 31-OCT-93.
DR PDB: 1IYU; 12-MAR-97.
DR PDB: 1IYV; 12-MAR-97.
DR PDB: 1DPB; 20-APR-95.
DR PDB: 1DPC; 20-APR-95.
DR PDB: 1DPD; 20-APR-95.
DR InterPro: IPR001078; 2Oxoacid_dh.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR004167; E3_binding.
DR InterPro: IPR003016; Lipoyl.
DR Pfam: PF00198; 2-oxoacid_dh; 1.
DR Pfam: PF00364; biotin_lipoyl; 3.
DR Pfam: PF02817; e3_binding; 1.
DR PRODOM: PD001115; 2Oxoacid_dh; 1.
DR PROSITE: PS00189; LIPOYL; 3.
KW Glycolysis; Transferase; Acyltransferase; Repeat; Lipoyl;
KW 3D-structure.
FT INIT_MET 0
FT DOMAIN 1 326 LIPOYL BINDING.
FT DOMAIN 327 380 E1/E3 BINDING.
FT DOMAIN 381 637 CATALYTIC.
FT BINDING 39 LIPOYL (POTENTIAL).
FT BINDING 156 LIPOYL (POTENTIAL).
FT BINDING 156 LIPOYL (POTENTIAL).
FT BINDING 261 LIPOYL (POTENTIAL).
FT REPEAT 1 115 LIPOYL (POTENTIAL).
FT REPEAT 116 220 LIPOYL (POTENTIAL).
FT REPEAT 221 326 LIPOYL (POTENTIAL).
FT ACT_SITE 610 610 POTENTIAL.
FT HELIX 403 406
FT HELIX 416 431

```



```
FT STRAND 434 442
FT HELIX 444 452
FT TURN 453 453
FT HELIX 454 459
FT TURN 460 461
FT HELIX 466 480
FT HELIX 482 484
FT STRAND 486 488
FT TURN 490 491
FT STRAND 495 497
FT STRAND 503 505
FT STRAND 507 509
FT TURN 510 511
FT STRAND 512 514
FT STRAND 517 518
FT HELIX 521 523
FT HELIX 526 541
FT TURN 542 543
FT HELIX 547 550
FT STRAND 555 559
FT TURN 561 563
FT TURN 574 575
FT STRAND 578 582
FT STRAND 586 591
FT STRAND 596 609
FT TURN 612 612
FT HELIX 615 630
FT HELIX 632 636
SQ SEQUENCE 637 AA; 64913 MW; D0603BA4A385F84 CRC64;
```

Query Match Best Local Similarity 7.1%; Score 144; DB 1; Length 637;  
Matches 80; Conservative 39; Mismatches 149; Indels 88; Gaps 11;

```
QY 19 PDASVADVKTITET-POGQSTYRADQOMLYOGKILDETTLESGVYENSFLYIMLSKA 77
DB 122 PDGSGAKAVIEVLVAKAGVOAEGSLVLESKASMEIIPSPASGVYES--VAIQUNAE 179
QY 78 KASSSGASTATKAPATLTAQPAAPVAPASVARTPTQAPVATATETAPPSVO---PQAA 133
DB 180 VGTGDDLTLFLRTTGA-----QAQPTAPAAAAASPAPAPLAPAAAGPOEVKVDIGSAG 233
QY 134 PAATVATDDADYYSQASNLVFGNNLEQITIQIILDMGGTWEEDTYVRALRAAYNNPER 193
DB 234 KAVIEVLVAKAGVOAEGSLI-----VLESDKASMEIIPSP 269
QY 194 AIDYISGIPENTYEAOPVARAPAPAGQOTNOQASPAQPAVALPYQSPASAGPANPLNL 253
DB 270 A-----AGVESYAVQUNAEVGTGDDILITLRVAGAAFSGRKASPFQAAAAAPAPAPA 324
QY 254 FPOGVPSGSGNPQVPGAGSALDALRQLPOFQALQLVOANPQIILQMLQELCKQNPQI 313
DB 325 -PYGAPSRN-----CAKVHAGPAVRQL-----AREGVE----- 352
QY 314 LRLIQEQAQAEFLKLVNESPESGPGNLTGOLAAAVPOTLVTPEPERAIOQLBSMG 369
DB 353 -----LAINST-----GPRGRLKEDVQAYVYKAMQKAEAPAGASGAG 394

RESULT 11
GSRL_HUMAN STANDARD; PRT; 1509 AA.
AC 09NMZ4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glioma tumor suppressor candidate region gene 1 protein.
GN GLTSCR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE-20175430; PubMed-10708517;
RA Smith J.S., Tachibana I., Pohl U., Lee H.K., Thanarajasingam U.,
RA Portier B.P., Deki K., Billings S., Ramaswamy S., Mohrenweiser H.W.,
RA Scheithauer B.W., Louis D.N., Jenkins R.B.;
RT "A transcript map of the chromosome 19q-Arm glioma tumor suppressor
region." 64:44-50(2000).
RL TISSUE SPECIFICITY: Expressed at moderate levels in heart, brain,
placenta, skeletal muscle, and pancreas, and at lower levels in
lung, liver, and kidney.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

Query Match Best Local Similarity 6.9%; Score 140.5; DB 1; Length 1509;  
Matches 108; Conservative 36; Mismatches 135; Indels 181; Gaps 24;

```
QY 55 DETTLESGVYA-ENSPFLYM--LSKAKASS-----SGASATPT-----AAPA 94
DB 220 EPTVLASAGVSPGAGATLQKNTLSAAVATTLNGSVFEGAGAAAPGTPSGOPLAVALPQ 279
QY 95 TLAQAPAVAPASVARTPTQAPVATATETAPPSV-----QQAAPATVATDDADYYSQ 149
DB 280 LGSSPLVP-APNVILHPTPTIQPKPAGVLPKLYQLPKFPAPAGATLITQSGRPLAQ 338
QY 150 ---AASNLVFGNNLEQITIQIILDMGGTWEEDTYVR---ALR-----AAY 188
DB 339 QPKAPQNLTF-----MAAGKAGQNVLSGFPAPALQANVFKQPPATTTGAAP 385
QY 189 NNPERAT-----DYLYSG-----IPENYEAQ----- 210
DB 386 PQQPAGLSKPSVHLNQGSSIVTPAQHMLPGQNFILPGAPAVQLPQQLSALPANYGQ 445
QY 211 --VARAPAGQO-----TNOQASP-----AQPAV-- 233
DB 446 ILAAAPHTGGQLANPLTNQNLAGPLSLCPVLAPHSGAHSAILLSAAPTQVQCPALPQ 505
QY 234 -----ALFPQSPASAGPNANP-----INLFPQGV-----PSGGSNPQVPGAGSAL 276
DB 506 MPVSLAAGSLPTQSQPACGPAATTVLQGVTLPPSAVAMLNTPGVLQVAPATPAATGEAA 565
QY 277 DALRQLPQFQALLQVQANPQIILQMLQELCKQNPQILRLIQEQAQAEFLRLV-----NSP 332
DB 566 PVLTVQ-----APQAPVAVSTPL--PLGLQDPQAO---QPPQAPTQQAAPAPQATTP 613
QY 333 EGGPG-----GNILGOLAAAVPOTLTV-----TPPER 359
DB 614 QPSPGLASSPEKTVLQGPSPATITAILITQDSLOMFLPQER 653

RESULT 12
CBP_MOUSE STANDARD; PRT; 2441 AA.
ID CBP_MOUSE
AC P45481;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
```

DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE CREB-binding protein.  
 GN CREBBP OR CBP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA MEDLINE=94019666; PubMed=8413673;  
 RA Christia J.C., Kwok R.P.S., Lamb N., Haglware M., Montminy M.R.,  
 RA Goodman R.H.;  
 RT "Phosphorylated CREB binds specifically to the nuclear protein CBP.";  
 RL Nature 365:855-859(1993).  
 CC -1- FUNCTION: MEDIATES CAMP-GENE REGULATION BY BINDING SPECIFICALLY TO  
 CC PHOSPHORYLATED CREB PROTEIN. ACTING AS A COACTIVATOR, CBP ARGUMENTS  
 CC THE ACTIVITY OF PHOSPHORYLATED CREB TO ACTIVATE TRANSCRIPTION OF  
 CC CAMP-RESPONSIVE GENES.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: CONTAINS 1 BROMODOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: S66385; AAB28651.1; -  
 DR TRANSFAC: T01318; -  
 DR MGD: MGI:1098280; Crebbp.  
 DR InterPro: IPR001487; Bromodomain.  
 DR InterPro: IPR003101; KIX.  
 DR InterPro: IPR000197; TAZ\_finger.  
 DR InterPro: IPR000433; ZnF\_ZZ.  
 DR Pfam: PF00439; bromodomain.1.  
 DR Pfam: PF00563; ZZ.1.  
 DR Pfam: PF02135; ZF\_TAZ.2.  
 DR Pfam: PF02172; KIX.1.  
 DR PRINTS: PRO0503; BROMODOMAIN.  
 DR SMART: SM00297; BROMO.1.  
 DR SMART: SM00291; ZnF\_ZZ.1.  
 DR PROSITE: PS00633; BROMODOMAIN.1; 1.  
 DR PROSITE: PS0014; BROMODOMAIN.2; 1.  
 DR PROSITE: PS01357; ZF\_ZZ.1; 1.  
 DR PROSITE: PS0135; ZF\_ZZ.2; 1.  
 KW Transcription regulation; Nuclear protein; Activator; Bromodomain;  
 KW Zinc-finger.  
 FT DOMAIN 1104 1176 BROMODOMAIN.  
 FT ZN\_FING 1702 1745 Z2-TYPE.  
 FT DOMAIN 1062 1065 POLY-GLU.  
 FT DOMAIN 1556 1563 POLY-GLU.  
 FT DOMAIN 1944 1949 POLY-PRO.  
 FT DOMAIN 1968 1971 POLY-GLN.  
 FT DOMAIN 2082 2086 POLY-GLN.  
 FT DOMAIN 2200 2216 POLY-GLN.  
 FT DOMAIN 2226 2299 POLY-GLN.  
 SQ SEQUENCE 2441 AA; 265474 MW; 0ABB028C3112F419 CRC64;  
 Query Match 6.9%; Score 140; DB 1; Length 2441;  
 Best Local Similarity 21.7%; Pred. No. 2.1;  
 Matches 92; Conservative 41; Mismatches 159; Indels 132; Gaps 18;  
 QY 44 QMLTYGKILDE-TTLESNGVAENSFLVIMLSKAKSSGASTATTAKAPATLAQPAAP 102  
 DB 157 QHCQAGQAMRRATMTATRVPOSL-----PSPISAPETPTQOSTPTQPP 1907  
 QY 103 VAPAS-----VART-PT---QAPVATETAPPSPVQAPAPATVATDAD 145  
 DB 1308 AQPSPSPVNMSPAGFPVNAFTQPTIYASGKPTQVNPAPPPAPD---PRAVEAARQIE 1964

QY 146 VYQAASNLVFGNNLEQTTIQIILDMGGGTWERDYVRLRAAYNNPERAIDYLSGIP- 204  
 DB 1965 REAQOQOHLXYRAN-----INNGMPG-----RDMGTPGSGMTFVGLNVRP 2006  
 QY 205 NVEADQVNAAPAPAGQOTN-----QQAASPAQPAVALPVQSPASAGPNAPLNFQGVPS 260  
 DB 2007 NOVSGFVMSMPGGWQQAIPQOQPMPCMPRVVSMQAAVAGPR-----MPN 2056  
 QY 261 GGSNPGVVGAGSGALDALROL-----PQFALLQLYVANPQIL----- 299  
 DB 2057 VQPNNSISF---SALQDLRLTKSPSSPQOQOQVNLILSNPQLMAAFIKQRTAYVANO 2113  
 QY 300 -----QPMQLQELGKONPQ-----ILRLIQNQAFLRLVNESPEGGPG-NILGQLAA 346  
 DB 2114 PCMQPQGLQSPGCMQPGCMHQPSLQVLMQAGVPRPGVPPQPMGSLNPGQALN 2173  
 QY 347 AV-----PQTLVTPEPERAQRRLBGMCFNRELVEVFACKNDELRANVYLDHGHEPD 401  
 DB 2174 INNPGHNPMPTMNPQYREMR-----QLLQHQOQOQ 2206  
 QY 402 PQOQ 405  
 DB 2207 QOQO 2210  
 RESULT 13  
 BAT3\_HUMAN STANDARD; PRT; 1132 AA.  
 ID BAT3\_HUMAN  
 AC P46379;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Large proline-rich protein BAT3 (HLA-B-associated transcript 3).  
 GN BAT3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=T-cell;  
 RA MEDLINE=90192810; PubMed=2156268;  
 RA Banerji J., Sands J., Strominger J.L., Spies T.;  
 RT "A gene pair from the human major histocompatibility complex encodes  
 RT large proline-rich proteins with multiple repeated motifs and a  
 RT single ubiquitin-like domain.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378(1990).  
 CC -1- FUNCTION: UNKNOWN.  
 CC -1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M33519; AAA35587.1; -  
 DR EMBL: M33521; AAA35588.1; -  
 DR EMBL: M33520; AAA35588.1; JOINED.  
 DR HSSP: P02248; UBI.  
 DR GeneW: HGNC:13919; BAT3.  
 DR MIM: 142590; -  
 DR InterPro: IPR000626; Ubiquitin.  
 DR Pfam: PF00240; ubiquitin.1.  
 DR SMART: SM00213; UBC.1.  
 DR PROSITE: PS00299; UBIQUITIN.1; 1.  
 DR PROSITE: PS0053; UBIQUITIN.2; 1.  
 KW Repeat.  
 FT DOMAIN 17 77 UBIQUITIN-LIKE.  
 FT DOMAIN 202 207 POLY-PRO.



